

OIPE

RAW SEQUENCE LISTING

DATE: 07/05/2001

PATENT APPLICATION: US/09/722,441

TIME: 16:42:45

Input Set : A:\1533 1030002 seq list.txt

Output Set : N:\CRF3\07032001\I722441.raw

5 <110> APPLICANT: Hanke, Paul D.
 7 Li-D'Elia, Lhing-Yew
 9 Rayapati, John
 11 Crafton, Corey
 13 Walsh, Holly
 17 <120> TITLE OF INVENTION: Increased Lysine Production by Gene Amplification
 21 <130> FILE REFERENCE: 1533.1030002
 25 <140> CURRENT APPLICATION NUMBER: 09/722,441
 C--> 27 <141> CURRENT FILING DATE: 2001-06-22
 30 <150> PRIOR APPLICATION NUMBER: US 60/173,707
 31 <151> PRIOR FILING DATE: 1999-12-30
 34 <150> PRIOR APPLICATION NUMBER: US 60/184,130
 35 <151> PRIOR FILING DATE: 2000-02-22
 39 <160> NUMBER OF SEQ ID NOS: 37
 43 <170> SOFTWARE: PatentIn version 3.0
 47 <210> SEQ ID NO: 1
 49 <211> LENGTH: 1266
 51 <212> TYPE: DNA
 53 <213> ORGANISM: Corynebacterium glutamicum
 57 <220> FEATURE:
 59 <221> NAME/KEY: CDS
 61 <222> LOCATION: (1)..(1266)
 65 <400> SEQUENCE: 1
 66 gtg gcc ctg gtc gta cag aaa tat ggc ggt tcc tcg ctt gag agt gcg 48
 67 Met Ala Leu Val Val Gln Lys Tyr Gly Gly Ser Ser Leu Glu Ser Ala
 68 1 5 10 15
 70 gaa cgc att aga aac gtc gct gaa cgg atc gtt gcc acc aag aag gct 96
 71 Glu Arg Ile Arg Asn Val Ala Glu Arg Ile Val Ala Thr Lys Lys Ala
 72 20 25 30
 74 gga aat gat gtc gtg gtt gtc tgc tcc gca atg gga gac acc acg gat 144
 75 Gly Asn Asp Val Val Val Val Cys Ser Ala Met Gly Asp Thr Thr Asp
 76 35 40 45
 78 gaa ctt cta gaa ctt gca gcg gca gtg aat ccc gtt ccg cca gct cgt 192
 79 Glu Leu Leu Glu Leu Ala Ala Ala Val Asn Pro Val Pro Pro Ala Arg
 80 50 55 60
 82 gaa atg gat atg ctc ctg act gct ggt gag cgt att tct aac gct ctc 240
 83 Glu Met Asp Met Leu Leu Thr Ala Gly Glu Arg Ile Ser Asn Ala Leu
 84 65 70 75 80
 86 gtc gcc atg gct att gag tcc ctt ggc gca gaa gct caa tct ttc act 288
 87 Val Ala Met Ala Ile Glu Ser Leu Gly Ala Glu Ala Gln Ser Phe Thr
 88 85 90 95
 90 ggc tct cag gct ggt gtg ctc acc acc gag cgc cac gga aac gca cgc 336
 91 Gly Ser Gln Ala Gly Val Leu Thr Thr Glu Arg His Gly Asn Ala Arg
 92 100 105 110
 94 att gtt gac gtc aca ccg ggt cgt gtg cgt gaa gca ctc gat gag ggc 384
 95 Ile Val Asp Val Thr Pro Gly Arg Val Arg Glu Ala Leu Asp Glu Gly
 96 115 120 125

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/722,441

DATE: 07/05/2001

TIME: 16:42:45

Input Set : A:\1533 1030002 seq list.txt

Output Set: N:\CRF3\07032001\I722441.raw

98	aag atc tgc att gtt gct ggt ttt cag ggt gtt aat aaa gaa acc cgc	432
99	Lys Ile Cys Ile Val Ala Gly Phe Gln Gly Val Asn Lys Glu Thr Arg	
100	130 135 140	
102	gat gtc acc acg ttg ggt cgt ggt ggt tct gac acc act gca gtt gcg	480
103	Asp Val Thr Thr Leu Gly Arg Gly Gly Ser Asp Thr Thr Ala Val Ala	
104	145 150 155 160	
106	ttg gca gct gct ttg aac gct gat gtg tgt gag att tac tcg gac gtt	528
107	Leu Ala Ala Ala Leu Asn Ala Asp Val Cys Glu Ile Tyr Ser Asp Val	
108	165 170 175	
110	gac ggt gtg tat acc gct gac ccg cgc atc gtt cct aat gca cag aag	576
111	Asp Gly Val Tyr Thr Ala Asp Pro Arg Ile Val Pro Asn Ala Gln Lys	
112	180 185 190	
114	ctg gaa aag ctc agc ttc gaa gaa atg ctg gaa ctt gct gct gtt ggc	624
115	Leu Glu Lys Leu Ser Phe Glu Glu Met Leu Glu Leu Ala Ala Val Gly	
116	195 200 205	
118	tcc aag att ttg gtg ctg cgc agt gtt gaa tac gct cgt gca ttc aat	672
119	Ser Lys Ile Leu Val Leu Arg Ser Val Glu Tyr Ala Arg Ala Phe Asn	
120	210 215 220	
122	gtg cca ctt cgc gta cgc tcg tct tat agt aat gat ccc ggc act ttg	720
123	Val Pro Leu Arg Val Arg Ser Ser Tyr Ser Asn Asp Pro Gly Thr Leu	
124	225 230 235 240	
126	att gcc ggc tct atg gag gat att cct gtg gaa gaa gca gtc ctt acc	768
127	Ile Ala Gly Ser Met Glu Asp Ile Pro Val Glu Glu Ala Val Leu Thr	
128	245 250 255	
130	ggt gtc gca acc gac aag tcc gaa gcc aaa gta acc gtt ctg ggt att	816
131	Gly Val Ala Thr Asp Lys Ser Glu Ala Lys Val Thr Val Leu Gly Ile	
132	260 265 270	
134	tcc gat aag cca ggc gag gct gcc aag gtt ttc cgt gcg ttg gct gat	864
135	Ser Asp Lys Pro Gly Glu Ala Ala Lys Val Phe Arg Ala Leu Ala Asp	
136	275 280 285	
138	gca gaa atc aac att gac atg gtt ctg cag aac gtc tcc tct gtg gaa	912
139	Ala Glu Ile Asn Ile Asp Met Val Leu Gln Asn Val Ser Ser Val Glu	
140	290 295 300	
142	gac ggc acc acc gac atc acg ttc acc tgc cct cgc gct gac gga cgc	960
143	Asp Gly Thr Thr Asp Ile Thr Phe Thr Cys Pro Arg Ala Asp Gly Arg	
144	305 310 315 320	
146	cgt gcg atg gag atc ttg aag aag ctt cag gtt cag ggc aac tgg acc	1008
147	Arg Ala Met Glu Ile Leu Lys Lys Leu Gln Val Gln Gly Asn Trp Thr	
148	325 330 335	
150	aat gtg ctt tac gac gac cag gtc ggc aaa gtc tcc ctc gtg ggt gct	1056
151	Asn Val Leu Tyr Asp Asp Gln Val Gly Lys Val Ser Leu Val Gly Ala	
152	340 345 350	
154	ggc atg aag tct cac cca ggt gtt acc gca gag ttc atg gaa gct ctg	1104
155	Gly Met Lys Ser His Pro Gly Val Thr Ala Glu Phe Met Glu Ala Leu	
156	355 360 365	
158	cgc gat gtc aac gtg aac atc gaa ttg att tcc atc tct gag atc cgc	1152
159	Arg Asp Val Asn Val Asn Ile Glu Leu Ile Ser Ile Ser Glu Ile Arg	
160	370 375 380	
162	att tcc gtg ctg atc cgt gaa gat gat ctg gat gct gct gca cgt gca	1200

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/722,441

DATE: 07/05/2001

TIME: 16:42:45

Input Set : A:\1533 1030002 seq list.txt

Output Set: N:\CRF3\07032001\I722441.raw

```

163 Ile Ser Val Leu Ile Arg Glu Asp Asp Leu Asp Ala Ala Ala Arg Ala
164 385                      390                      395                      400
166 ttg cat gag cag ttc cag ctg ggc ggc gaa gac gaa gcc gtc gtt tat      1248
167 Leu His Glu Gln Phe Gln Leu Gly Gly Glu Asp Glu Ala Val Val Tyr
168                      405                      410                      415
170 gca ggc acc gga cgc taa      1266
171 Ala Gly Thr Gly Arg
172                      420
175 <210> SEQ ID NO: 2
177 <211> LENGTH: 421
179 <212> TYPE: PRT
181 <213> ORGANISM: Corynebacterium glutamicum
185 <400> SEQUENCE: 2
187 Met Ala Leu Val Val Gln Lys Tyr Gly Gly Ser Ser Leu Glu Ser Ala
188 1                      5                      10                      15
191 Glu Arg Ile Arg Asn Val Ala Glu Arg Ile Val Ala Thr Lys Lys Ala
192                      20                      25                      30
195 Gly Asn Asp Val Val Val Val Cys Ser Ala Met Gly Asp Thr Thr Asp
196                      35                      40                      45
199 Glu Leu Leu Glu Leu Ala Ala Val Asn Pro Val Pro Pro Ala Arg
200 50                      55                      60
203 Glu Met Asp Met Leu Leu Thr Ala Gly Glu Arg Ile Ser Asn Ala Leu
204 65                      70                      75                      80
207 Val Ala Met Ala Ile Glu Ser Leu Gly Ala Glu Ala Gln Ser Phe Thr
208                      85                      90                      95
211 Gly Ser Gln Ala Gly Val Leu Thr Thr Glu Arg His Gly Asn Ala Arg
212                      100                     105                     110
215 Ile Val Asp Val Thr Pro Gly Arg Val Arg Glu Ala Leu Asp Glu Gly
216                      115                     120                     125
219 Lys Ile Cys Ile Val Ala Gly Phe Gln Gly Val Asn Lys Glu Thr Arg
220                      130                     135                     140
223 Asp Val Thr Thr Leu Gly Arg Gly Gly Ser Asp Thr Thr Ala Val Ala
224 145                     150                     155                     160
227 Leu Ala Ala Ala Leu Asn Ala Asp Val Cys Glu Ile Tyr Ser Asp Val
228                      165                     170                     175
231 Asp Gly Val Tyr Thr Ala Asp Pro Arg Ile Val Pro Asn Ala Gln Lys
232                      180                     185                     190
235 Leu Glu Lys Leu Ser Phe Glu Glu Met Leu Glu Leu Ala Ala Val Gly
236                      195                     200                     205
239 Ser Lys Ile Leu Val Leu Arg Ser Val Glu Tyr Ala Arg Ala Phe Asn
240                      210                     215                     220
243 Val Pro Leu Arg Val Arg Ser Ser Tyr Ser Asn Asp Pro Gly Thr Leu
244 225                     230                     235                     240
247 Ile Ala Gly Ser Met Glu Asp Ile Pro Val Glu Glu Ala Val Leu Thr
248                      245                     250                     255
251 Gly Val Ala Thr Asp Lys Ser Glu Ala Lys Val Thr Val Leu Gly Ile
252                      260                     265                     270
255 Ser Asp Lys Pro Gly Glu Ala Ala Lys Val Phe Arg Ala Leu Ala Asp
256                      275                     280                     285

```

RAW SEQUENCE LISTING

DATE: 07/05/2001

PATENT APPLICATION: US/09/722,441

TIME: 16:42:45

Input Set : A:\1533 1030002 seq list.txt

Output Set: N:\CRF3\07032001\I722441.raw

```

259 Ala Glu Ile Asn Ile Asp Met Val Leu Gln Asn Val Ser Ser Val Glu
260      290      295      300
263 Asp Gly Thr Thr Asp Ile Thr Phe Thr Cys Pro Arg Ala Asp Gly Arg
264 305      310      315      320
267 Arg Ala Met Glu Ile Leu Lys Lys Leu Gln Val Gln Gly Asn Trp Thr
268      325      330      335
271 Asn Val Leu Tyr Asp Asp Gln Val Gly Lys Val Ser Leu Val Gly Ala
272      340      345      350
275 Gly Met Lys Ser His Pro Gly Val Thr Ala Glu Phe Met Glu Ala Leu
276      355      360      365
279 Arg Asp Val Asn Val Asn Ile Glu Leu Ile Ser Ile Ser Glu Ile Arg
280      370      375      380
283 Ile Ser Val Leu Ile Arg Glu Asp Asp Leu Asp Ala Ala Ala Arg Ala
284 385      390      395      400
287 Leu His Glu Gln Phe Gln Leu Gly Gly Glu Asp Glu Ala Val Val Tyr
288      405      410      415
291 Ala Gly Thr Gly Arg
292      420
295 <210> SEQ ID NO: 3
297 <211> LENGTH: 1035
299 <212> TYPE: DNA
301 <213> ORGANISM: Corynebacterium glutamicum
305 <220> FEATURE:
307 <221> NAME/KEY: CDS
309 <222> LOCATION: (1)..(1035)
313 <400> SEQUENCE: 3
314 atg acc acc atc gca gtt gtt ggt gca acc ggc cag gtc ggc cag gtt      48
315 Met Thr Thr Ile Ala Val Val Gly Ala Thr Gly Gln Val Gly Gln Val
316 1      5      10      15
318 atg cgc acc ttt ttg gaa gag cgc aat ttc cca gct gac act gtt cgt      96
319 Met Arg Thr Phe Leu Glu Glu Arg Asn Phe Pro Ala Asp Thr Val Arg
320      20      25      30
322 ttc ttt gct tcc ccg cgt tcc gca ggc cgt aag att gaa ttc cgt ggc      144
323 Phe Phe Ala Ser Pro Arg Ser Ala Gly Arg Lys Ile Glu Phe Arg Gly
324      35      40      45
326 acg gaa atc gag gta gaa gac att act cag gca acc gag gag tcc ctc      192
327 Thr Glu Ile Glu Val Glu Asp Ile Thr Gln Ala Thr Glu Glu Ser Leu
328      50      55      60
330 aag ggc atc gac gtt gcg ttg ttc tct gct gga ggc acc gct tcc aag      240
331 Lys Gly Ile Asp Val Ala Leu Phe Ser Ala Gly Gly Thr Ala Ser Lys
332 65      70      75      80
334 cag tac gct cca ctg ttt gct gct gca ggc gcg act gtt gtg gat aac      288
335 Gln Tyr Ala Pro Leu Phe Ala Ala Ala Gly Ala Thr Val Val Asp Asn
336      85      90      95
338 tct tct gct tgg cgc aag gac gac gag gtt cca cta atc gtc tct gag      336
339 Ser Ser Ala Trp Arg Lys Asp Asp Glu Val Pro Leu Ile Val Ser Glu
340      100      105      110
342 gtg aac cct tcc gac aag gat tcc ctg gtc aag ggc att att gcg aat      384
343 Val Asn Pro Ser Asp Lys Asp Ser Leu Val Lys Gly Ile Ile Ala Asn

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/722,441

DATE: 07/05/2001

TIME: 16:42:45

Input Set : A:\1533 1030002 seq list.txt

Output Set: N:\CRF3\07032001\I722441.raw

```

344          115          120          125
346 cct aac tgc acc acc atg gct gca atg cca gtg ctg aag cca ctg cac      432
347 Pro Asn Cys Thr Thr Met Ala Ala Met Pro Val Leu Lys Pro Leu His
348          130          135          140
350 gat gcc gct ggt ctt gta aag ctt cac gtt tcc tct tac cag gct gtt      480
351 Asp Ala Ala Gly Leu Val Lys Leu His Val Ser Ser Tyr Gln Ala Val
352 145          150          155          160
354 tcc ggt tct ggt ctt gca ggt gtg gaa acc ttg gca aag cag gtt gct      528
355 Ser Gly Ser Gly Leu Ala Gly Val Glu Thr Leu Ala Lys Gln Val Ala
356          165          170          175
358 gca gtt ggc gac cac aac gtt gag ttc gtc cat gat gga cag gct gct      576
359 Ala Val Gly Asp His Asn Val Glu Phe Val His Asp Gly Gln Ala Ala
360          180          185          190
362 gac gca ggc gat gtc gga cct tac gtt tcc cca atc gct tac aac gtg      624
363 Asp Ala Gly Asp Val Gly Pro Tyr Val Ser Pro Ile Ala Tyr Asn Val
364          195          200          205
366 ctg cca ttc gcc gga aac ctc gtc gat gac ggc acc ttc gaa acc gac      672
367 Leu Pro Phe Ala Gly Asn Leu Val Asp Asp Gly Thr Phe Glu Thr Asp
368          210          215          220
370 gaa gag cag aag ctg cgc aac gaa tcc cgc aag att ctc ggc ctc cca      720
371 Glu Glu Gln Lys Leu Arg Asn Glu Ser Arg Lys Ile Leu Gly Leu Pro
372 225          230          235          240
374 gac ctc aag gtc tca ggc acc tgc gtc cgc gtg ccg gtt ttc acc ggc      768
375 Asp Leu Lys Val Ser Gly Thr Cys Val Arg Val Pro Val Phe Thr Gly
376          245          250          255
378 cac acg ctg acc att cac gcc gaa ttc gac aag gca atc acc gtc gag      816
379 His Thr Leu Thr Ile His Ala Glu Phe Asp Lys Ala Ile Thr Val Glu
380          260          265          270
382 cag gcg cag gag atc ttg ggt gcc gct tca ggc gtc gag ctt gtc gac      864
383 Gln Ala Gln Glu Ile Leu Gly Ala Ala Ser Gly Val Glu Leu Val Asp
384          275          280          285
386 gtc cca acc cca ctt gca gct gcc ggc att gac gaa tcc ctc gtt gga      912
387 Val Pro Thr Pro Leu Ala Ala Ala Gly Ile Asp Glu Ser Leu Val Gly
388          290          295          300
390 cgc atc cgt cag gac tcc act gtc gac gac aac cgc ggt ctg gtt ctc      960
391 Arg Ile Arg Gln Asp Ser Thr Val Asp Asp Asn Arg Gly Leu Val Leu
392 305          310          315          320
394 gtc gta tct ggc gat aac ctt cgc aag ggc gca gca ctg aac acc att      1008
395 Val Val Ser Gly Asp Asn Leu Arg Lys Gly Ala Ala Leu Asn Thr Ile
396          325          330          335
398 cag att gct gag ctg ctg gtt aag taa      1035
399 Gln Ile Ala Glu Leu Leu Val Lys
400          340
403 <210> SEQ ID NO: 4
405 <211> LENGTH: 344
407 <212> TYPE: PRT
409 <213> ORGANISM: Corynebacterium glutamicum
413 <400> SEQUENCE: 4
415 Met Thr Thr Ile Ala Val Val Gly Ala Thr Gly Gln Val Gly Gln Val

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/722,441

DATE: 07/05/2001

TIME: 16:42:46

Input Set : A:\1533 1030002 seq list.txt

Output Set: N:\CRF3\07032001\I722441.raw

L:27 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:2037 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:22
L:2055 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:23
L:2073 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:24
L:2091 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:25
L:2109 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:26
L:2127 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:27
L:2145 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:28
L:2163 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:29
L:2181 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:30
L:2199 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:31
L:2217 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:32
L:2235 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:33
L:2253 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:34